



BEILAGE ZUR FACHZEITSCHRIFT SOZIALE SICHERHEIT
ERSTELLT DURCH DAS INSTITUT FÜR HÖHERE STUDIEN, IHS HealthEcon
Herausgegeben vom Hauptverband der österreichischen Sozialversicherungsträger

Health System Watch



Performance von Gesundheitssystemen

Thomas Czypionka, Gerald Röhrling, Marcel Kalmar*

Zusammenfassung

Performance-Messung von Gesundheitssystemen dient primär dazu, Informationen für alle im Gesundheitssystem beteiligten Stakeholder zur Verfügung zu stellen; somit leistet dieses Instrument einen wichtigen Beitrag zur Verbesserung der Transparenz im Gesundheitssystem. Im Zuge der Performance-Messung müssen Zielvorstellungen, was Gesundheitssysteme leisten sollen, genau formuliert und operationalisiert sowie Bereiche mit passenden Indikatoren ausgewählt werden, welche die dahinterstehenden Konstrukte auch tatsächlich abbilden. Bisherige Versuche internationaler Performance-Messungen von Gesundheitssystemen – allen voran der World Health Report 2000 der WHO – wurden oftmals starker Kritik unterzogen, insbesondere dahingehend, dass die Messungen für zu viele unterschiedliche Länder mit schlechten Indikatoren basierend auf inhomogenen und nicht nachvollziehbaren Datengrundlagen durchgeführt wurden. Daher sollte jedes publizierte Gesundheits-(system)-Ranking einer genauen und kritischen Analyse unterzogen werden. Trotzdem stellt eine wissenschaftlich fundierte Performance-Messung als Informationsinstrument ein anstrebenwertes Ziel dar, um auch den politischen Entscheidungsträgern zu helfen, Politiken zu formulieren, zu evaluieren oder als Regulativ einzugreifen. Dazu sollten Performance-Messungen aber nicht auf Gesamtrankings reduziert werden, sondern insbesondere eine höhere Informationsdichte und -komplexität aufweisen.

Einleitung

Die Messung der Leistungsfähigkeit (oder Performance) eines Gesundheitssystems sowie deren Einordnung in einen internationalen Kontext stellen von jeher eine Herausforderung an Wissenschaft und Forschung dar. Gründe dafür finden sich beispielsweise im hohen Komplexitätsgrad von Gesundheitssystemen sowie der Vielzahl von länderspezifisch verschieden ausgestalteten und einflussreichen Stakeholdern. Erste Entwicklungen zur Performance-Messung in einem Gesundheitssystem gingen bereits in den 1860ern von Florence Nightingale, einer britischen Krankenschwester und Statistikerin, aus, die eine systematische Erfassung von Krankenhausdaten, deren Analyse und Verbreitung initiierte.¹ Gesundheitssysteme in ihrer Gesamtheit zu bewerten ist erst in der jüngeren Vergangenheit durch die moderne Informationsverarbeitung bzw. Datengenerierung möglich geworden. Dabei verfolgen die internationalen Performance-Messungen von Gesundheitssystemen meist die Intention, das Erreichen spezifischer Gesundheitsoutcomes mit verschiedenen zugrunde liegenden Leistungserbringungs- und Finanzierungsmechanismen des Gesundheitswesens zu vergleichen. In diesem Zusammenhang ist es wichtig, im Hinterkopf zu behalten, dass die Performance eines Gesundheitssystems nicht notwendigerweise mit dem Gesundheitszustand der Bevölkerung korreliert sein muss, da dieser von weit mehr Einflussfaktoren als dem Gesundheitssystem determiniert wird – Genetik, sozioökonomischer Status, Politikmaßnahmen, Bildung etc.²

* Alle: Institut für Höhere Studien
Stumpergasse 56, A-1060 Wien, Telefon: +43/1/599 91-127, E-Mail: thomas.czypionka@ihs.ac.at.
Frühere Ausgaben von Health System Watch sind abrufbar im Internet unter: <http://www.ihs.ac.at>.

1 Smith P. C. (2002): Measuring health system performance. *European Journal of Health Economics*, 3, S. 145.
2 Canadian Institute for Health Information (CIHI) (2008): Making Sense of Health Rankings. Ottawa, S. 4.

Zusammenfassung

Einleitung





Vor mittlerweile rund einer Dekade versuchte die Weltgesundheitsorganisation (WHO) erstmals in ihrem jährlich publizierten World Health Report³ die Leistungsfähigkeit von Gesundheitssystemen weltweit zu messen. Dieser Bericht, der als finales Ergebnis ein Ranking aller 191 WHO-Mitgliedsstaaten präsentierte, löste nicht nur medial, sondern auch in der wissenschaftlichen Community eine rege Diskussion aus. Insbesondere Medien und politische Entscheidungsträger nahmen die Studie zum Anlass, auf vermeintliche Defizite ihres jeweiligen nationalen Gesundheitssystems hinzuweisen oder dieses in den Himmel zu heben. Die Kritik vieler Experten und Gesundheitssystemforscher ließ jedoch nicht lange auf sich warten. So wurde insbesondere kritisiert, dass es nicht möglich und sinnvoll ist, Gesundheitssysteme von 191 politisch, gesellschaftlich und ökonomisch verschiedenen Ländern miteinander zu vergleichen.⁴ Zudem wurden die Zuverlässigkeit der zugrunde liegenden Basisdaten vieler Länder, die unterschiedlichen Erhebungs- und Bezugszeitpunkte sowie die Auswahl und Gewichtung der Indikatoren heftiger Kritik unterzogen. Bis dato wurde seitens der WHO kein Update des 2000 publizierten Rankings, aber auch kein konzeptionell neuer Versuch, die Leistungsfähigkeit von Gesundheitssystemen in Form eines Rankings abzubilden, präsentiert. Nach einigen Jahren der „Stille“ publizierte das Health Consumer Powerhouse (HCP)⁵ im Jahre 2005 erstmals wieder ein internationales Ranking für damals zwölf europäische Gesundheitssysteme, den Euro Health Consumer Index (EHCI). Seit diesem Zeitpunkt veröffentlicht das Unternehmen jährlich – in zum Teil recht abgeänderter Form – für mittlerweile 33 Länder ein Ranking. Bemühungen zum Vergleich der Leistungsfähigkeit von Gesundheitssystemen haben aber auch bei der Europäischen Ministerkonferenz⁶ in Tallinn im Juni 2008 neues Interesse erfahren.

Ziel dieser Ausgabe des Health System Watch ist es, einen Überblick über verschiedene Ansätze zu geben, wie die Leistungsfähigkeit von Gesundheitssystemen gemessen werden kann, welche Bereiche abgedeckt werden (sollten) bzw. welche Konzepte zur Performance-Messung welche Vor- und Nachteile mit sich bringen. Weiters findet eine kritische Auseinandersetzung mit bereits bestehenden Gesundheitsrankings statt, und wir wollen auch einige Erkenntnisse wiedergeben, auf welche Aspekte bei der Interpretation derartiger Darstellungen und Berechnungen besonders achtgegeben werden sollte.

Wozu Performance-Messungen? Die Nutzenfrage

In erster Linie dient die Performance-Messung von Gesundheitssystemen der Verfügbarkeit von Informationen für alle Stakeholder. Anbieter erhalten Informationen über deren Stärken und Schwächen, Finanziere darüber, wie effizient ihr Geld verwendet wird, und Konsumenten erhalten u. a. Aufschluss über die Qualität ihrer medizinischen Versorgung. Die so geschaffene Transparenz soll aber keinem Selbstzweck dienen. Verantwortlichkeiten von Leistungsanbietern gegenüber ihren Patienten und von Regierungen gegenüber ihren Bürgern bzw. Patienten-Empowerment durch gezielte Information sind wesentliche Funktionen, welche Performance-Messungen erfüllen sollen.⁷ Dabei muss klar sein, dass abhängig von der Position in der Stakeholder-Struktur unterschiedliche Perspektiven auf das Gesundheitssystem eingenommen werden können und verschiedene Informationsbedürfnisse vorhanden sind. So benötigen Patienten für die Wahl eines Anbieters wesentlich detailliertere Informationen als ein Wähler, der die Arbeit seiner Regierung bzgl. der Gesundheitsversorgung beurteilen will. Eine weitere essenzielle Aufgabe der Performance-Messungen ist das Eruiere von „Best practice“-Strategien. Diese unterstützen politische Entscheidungsträger in ihrer Arbeit zur Verbesserung des Gesamtsystems und geben Anbietern die Möglichkeit, ungenutzte Potenziale ausfindig zu machen.⁸ Das bedeutet aber keineswegs, dass das

Wozu
Performance-
Messungen?
Die Nutzenfrage

3 WHO (2000): The World Health Report 2000. Health Systems: Improving Performance. Genf.

4 Beske F.: Das Gesundheitswesen in Deutschland im internationalen Vergleich. Eine Antwort auf die Kritik. Schriftenreihe des IGSF, Band 100. Kiel 2004.

McKee M.: The World Health Report 2000: 10 years on. Health Policy Plan. 2010 Sep;25(5):346-8.

5 HCP ist ein privates schwedisches Unternehmen mit Sitzen in Stockholm, Brüssel und Winnipeg, Kanada; die Finanzierung erfolgt über Forschungsgelder von Stakeholdern sowie über eine Kooperation mit der Europäischen Kommission.

6 Vgl. z. B. Europäische Ministerkonferenz der WHO zum Thema Gesundheitssysteme: Gesundheitssysteme, Gesundheit und Wohlstand: Die Charta von Tallinn: Gesundheitssysteme für Gesundheit und Wohlstand, Juni 2008.

7 Smith P. C., Mossialos E., Papanicolas I. (2008): Leistungsmessung zur Verbesserung der Gesundheitssysteme. Erfahrungen, Herausforderungen und Aussichten. Hintergrundpapier. Kopenhagen. WHO-Regionalbüro Europa, S.1.

8 Naylor D. C., Iron K., Handa K. (2002): Measuring Health System Performance: Problems And Opportunities In The Era Of Assessment And Accountability. In: Smith P. C. (Hrsg.) Measuring Up. Improving Health System Performance In OECD Countries. Paris. OECD, S. 14.





Kopieren von leistungsfähigen (Teil-)Systemen und Gesundheitspraxen zum Erfolg führen muss. Länderspezifische Anpassungen und Übersetzungen sind unumgänglich und führen realistischerweise nur zu Umsetzungen in kleinen Schritten.⁹

Ein oft diskutierter Punkt neben dem Nutzen sind die Qualität und Aussagekraft einer Performance-Messung. Probleme in der Datenverfügbarkeit und/oder deren Validität, Zurechnungsprobleme und sonstige methodische Schwierigkeiten sowie die Kosten der Untersuchungen sind einige wichtige Aspekte, die zu Skepsis gegenüber Performance-Messungen von Gesundheitssystemen führen.¹⁰

Konzeptionelle Entscheidungen im Zuge von Performance-Messungen

Ausgangspunkt für die Messung der Leistungsfähigkeit von Gesundheitssystemen sind deren Ziele. Das bedeutet, es muss geklärt werden, was Gesundheitssysteme überhaupt leisten sollen. Im Allgemeinen lassen sich trotz unterschiedlicher historischer und soziokultureller Entwicklungen (zumindest für OECD-Länder) einige wenige grob formulierte Ziele – wie beispielsweise die bedarfs- und chancengerechte, effiziente und effektive Versorgung mit Gesundheitsdienstleistungen – hervorheben.¹¹ Die Frage, die sich in diesem Zusammenhang stellt, ist, wie diese Zielvorstellungen in eine Performance-Messung übersetzt werden.

Systemgrenzen und Bereichsauswahl

Der naheliegendste Ansatz, die Performance eines Gesundheitssystems zu messen, ist die Gesundheit an sich. Dabei stellt sich die Frage, inwieweit das Gesundheitssystem überhaupt für die allgemeine Gesundheit verantwortlich ist. Wie bereits erwähnt beeinflussen vielerlei Faktoren den Gesundheitszustand der Bevölkerung, wobei diese weder im Gesundheitssystem entstehen noch in diesem gemanagt werden können. Theoretisch kann daher ein Land ein ausgezeichnetes Gesundheitssystem haben, obwohl Analysen des Gesundheitszustandes der Bevölkerung ein schlechtes Bild – relativ zu anderen Staaten – ergeben.¹² Ein Beispiel dafür ist die Lebenserwartung in Dänemark, die um ein Jahr geringer ist als in Österreich. Die erste konzeptionelle Entscheidung betrifft somit die Grenzen des Einflusses des Gesundheitssystems auf den Gesundheitszustand der Bevölkerung. Eine zusätzliche Erschwernis bringt die ambivalente Auslegung des Begriffs „Gesundheitssystem“ mit sich. So inkludiert die WHO im World Health Report 2000 jegliche Aktivität, die sich der Verbesserung, Wiederherstellung oder Beibehaltung des Gesundheitszustands widmet, in den Begriff des Gesundheitssystems. Die Wahl eines so breiten Gesundheitssystem-Begriffs erscheint nachvollziehbar, wenn unterschiedliche Länder mit enormen Systemunterschieden parallel untersucht werden. Dieser breite Systembegriff macht allerdings die Operationalisierung schwieriger, da administrative Daten nicht nach dieser Abgrenzung erfasst werden. Andere Performance-Messungen hingegen beschränken sich beispielsweise auf Definitionen, die ausschließlich eine staatlich anerkannte, professionelle Erbringung von Gesundheitsdienstleistungen umfassen.

Neben der Bevölkerungsgesundheit an sich gibt es weitere Aspekte und Bereichen, die als notwendige Bestandteile eines funktionsfähigen Gesundheitssystems aufgefasst werden, vgl. Tabelle 1. Dazu können beispielsweise die Bereiche Bedarfsgerechtigkeit eines Systems, die klinische Qualität und Angemessenheit der Versorgung, Chancengleichheit oder Effizienz des Systems gezählt werden. Für jeden Bereich stehen nun verschiedene Indikatoren(bündel) zur Verfügung, welche die Leistungsfähigkeit der Gesundheitssysteme quantifizieren sollen. Ob nur einzelne oder mehrere Bereiche in die Performance-Messung einfließen, determiniert die Perspektive und somit die Nutzbarkeit der Informationen für die jeweiligen Stakeholder. Je mehr Bereiche abgedeckt werden können, desto höher ist die Anzahl an Stakeholder-Gruppen, die Nutzen aus der Messung ziehen können. Dementsprechend steigt aber auch der finanzielle und wissenschaftliche Aufwand.

9 Naylor D. C., Iron K., Handa K. (2002): op. cit., S. 14.

10 vgl. Loeb J. M. (2004): The current state of performance measurement in health care. International Journal for Quality in Health Care, 16(1), S. i5.

11 Naylor D. C., Iron K., Handa K. (2002): op. cit., S. 14.

12 Canadian Institute for Health Information (CIHI) (2008): op.cit., S. 7.

**Konzeptionelle
Entscheidungen
im Zuge von
Performance-
Messungen**





Tabelle 1: Bereiche und ausgewählte Indikatoren einer Performance-Messung von Gesundheitssystemen

Bereich der Messung	Beschreibung des Maßes	Beispiele für Indikatoren
Bevölkerungs-gesundheit	Aggregierte Daten über die Gesundheit der Bevölkerung	Lebenserwartung Verlorene Lebensjahre vermeidbare Sterblichkeit DALY (um Behinderungen bereinigte Lebensjahre)
Gesundheits-ergebnisse für den Einzelnen	Gesundheitszustand einzelner Personen im Verhältnis zu einer Bevölkerung oder Gruppe Ranglisten von Gesundheitseinrichtungen nach Gesundheitszustand	Generische Indikatoren: ♦ SF-36 ^a ♦ EQ-5D ^b Krankheitsbezogene Indikatoren: ♦ Skala für Auswirkungen von Arthritis ♦ Parkinson-Fragebogen (PDQ-39)
Klinische Qualität und Angemessenheit der Versorgung	Angebote und Versorgung für Patienten zur Erreichung der gewünschten Ergebnisse Überprüfung auf bestbewährte Praktiken und fachlich vernünftige Anwendung	Ergebnisindikatoren: ♦ Gesundheitszustand ♦ konkrete Wiederaufnahme- und Sterblichkeitsraten nach Operationen Verfahrensindikatoren: ♦ Häufigkeit des Blutdruckmessens
Bedarfs-gerechtigkeit des Gesundheitssystems	Umgang mit den Menschen und Umfeld ihrer Behandlung im Gesundheitssystem Patientenwürde, Autonomie, Schweigepflicht, Kommunikation, prompte Bedienung, soziale Wärme und grundlegende Höflichkeit	Indikatoren für Patientenerfahrungen Indikatoren für Patientenzufriedenheit
Chancen-gleichheit	Ausmaß der Chancengleichheit in Bezug auf Gesundheit, Gesundheitsversorgung, Bedarfsgerechtigkeit und Finanzierung	Nutzungsindikatoren Zugangsraten Verhältnis Nutzung/Bedarf Ausgabenschwellen Disaggregierte Indikatoren gesundheitlicher Ergebnisse
Produktivität	Produktivität des Gesundheitssystems, der Versorgungsorganisationen und der einzelnen Ärzte	Arbeitsproduktivität Kostenwirksamkeitsindikatoren (für Interventionen) Fachliche Effizienz (Output/Input) Allokative Effizienz (gemessen an der Zahlungswilligkeit)

a SF-36 ist ein kurzer Mehrzweckfragebogen über Gesundheit mit nur 36 Fragen.

b EQ-5D ist ein Standardinstrument zur Messung von Behandlungsergebnissen für eine breite Auswahl von Erkrankungen. Damit wird ein einfaches deskriptives Profil und ein einzelner Indexwert für den Gesundheitszustand gewonnen, der für die klinische und ökonomische Auswertung der Gesundheitsversorgung sowie für Gesundheitssurveys benutzt werden kann.

Quelle: Smith P. C., Mossialos E., Papanicolas I., Leatherman S. (2010): Performance Measurement for Health System Improvement. Experiences, Challenges and Prospects. Cambridge. Cambridge University Press.

Indikatorenauswahl

Nach der Auswahl von Bereichen ist natürlich von entscheidender Bedeutung, durch welche Indikatoren diese abgebildet werden sollen. Die Güte einer Performance-Messung ist insbesondere von dieser Indikatorenauswahl sowie der Qualität der dahinterstehenden Daten abhängig. Bei der Datenauswahl können zwei Strategien verfolgt werden: Einerseits können bereits vorhandene Datensätze aus verschiedenen Gesundheitsbereichen eingesetzt werden. Diese Daten können meist kosteneffizient bezogen und verwendet werden, jedoch entfällt die konzeptionelle Komponente der gezielten Indikatorenauswahl, bzw. es stellt sich die Frage, wie relevant die verfügbaren Daten für die Performance-Messung überhaupt sind. Andererseits gibt es die Möglichkeit – einem konzeptionellen Rahmen folgend – neue Daten zur Indikatorenbildung zu generieren. Damit sind aber meist hoher Aufwand und Kosten verbunden, die mitunter die zusätzlich möglichen Informationen nicht aufwiegen.¹³

Gesundheitszustand

Indikatoren zur Feststellung der Bevölkerungsgesundheit wie etwa die allgemeine Lebenserwartung oder vorzeitig verlorene Lebensjahre sind weitverbreitete Maße zur Bewertung des allgemeinen Gesundheitszustands. Auf der anderen Seite ist der Grad des Einflusses des Gesundheitssystems auf diese Kennzahlen unbekannt. Dies gilt genauso auch für differenziertere Indikatoren wie





Disability Adjusted Life Years (DALYs) oder Disability Adjusted Life Expectancy (DALE), die Behinderungen (je nach Schweregrad) in die Berechnungen von verlorenen Lebensjahren oder der Lebenserwartung einbeziehen. J. P. Bunker¹⁴ untersuchte in den Vereinigten Staaten die Auswirkungen von mehreren präventiven und therapeutischen Maßnahmen auf die Lebenserwartung. Dabei wurden Behandlungen ausgewählt, die einerseits gute klinische Ergebnisse liefern und andererseits Krankheiten und Gesundheitsrisiken, welche eine ausreichend hohe Prävalenz aufweisen, dass sie für den Gesundheitszustand der Gesamtbevölkerung relevant sind. Bunker beziffert die gewonnenen Jahre an Lebenserwartung durch die gewählten medizinischen Interventionen seit 1900 auf 7 bis 7,5 Jahre, während die durchschnittliche Lebenserwartung um zirka 30 Jahre stieg. Für den Zeitraum 1950 bis 2000 hingegen schätzt Bunker die Bedeutung medizinischer Maßnahmen für die Lebenserwartung auf bis zu 50 Prozent.

Der Einfluss des Gesundheitssystems insgesamt auf die Bevölkerungsgesundheit bleibt bis heute nicht klar eruierbar. In diesem Zusammenhang sei aber auch auf neuere Konzepte wie etwa das der „vermeidbaren Sterblichkeit“ hingewiesen. Hierbei handelt es sich um die Auswertung von Sterbefällen, die aufgrund von existierenden medizinischen Interventionen hätten verhindert werden können. Im Allgemeinen lässt sich feststellen, dass klinische Ergebnisse am besten geeignet sind, um aussagekräftige Indikatoren für den Effekt des Gesundheitssystems auf den Gesundheitszustand zu entwickeln.¹⁵ Damit ist aber noch keine Entscheidung getroffen über die Auswahl, welche Interventionen für die Performance-Messung herangezogen werden sollen und anhand welcher Kriterien die Behandlungsergebnisse als positiv zu bewerten sind. Letzteres gilt vor allem für chronische oder psychische Erkrankungen, deren Erfolg äußerst schwer zu messen ist.

Bedarfsgerechtigkeit (Responsiveness)

Bedarfsgerechtigkeit umfasst die Art und Weise, wie mit dem Patienten im Gesundheitssystem verfahren wird. Dies inkludiert u. a. die Patientenwürde, die sich einerseits auf fundamentale Menschenrechte bezieht und etwa das Verbot von rassistischen oder religionsbedingten Diskriminierungen im Gesundheitssystem umfasst. Andererseits handelt es sich um die Betonung des Respekts vor der Person, die adäquates Verhalten gegenüber dem Patienten und noch wichtiger die Privatsphäre und die Autonomie des Patienten anerkennt.¹⁶ Neben den Persönlichkeitsrechten bezieht sich Responsiveness von Gesundheitssystemen oft auf die Kundenorientierung des Systems. Darunter fallen beispielsweise nicht-gesundheitliche Aspekte wie soziale Betreuung, Wartezeiten oder Wahlfreiheiten bezüglich Gesundheitsangeboten. Im Gegensatz zu Gesundheitsergebnissen repräsentiert Bedarfsgerechtigkeit die subjektiven Einschätzungen und Erfahrungen von Personen über das Gesundheitssystem, eben aus Konsumentenperspektive. Problematisch bei Indikatoren zur Messung der Responsiveness ist die hohe Abhängigkeit von sozialen und kulturellen Unterschieden. Da die Bezugsgrößen, mit denen die Bedarfsgerechtigkeit eines Gesundheitssystems gemessen wird, offen und mannigfaltig sind, bleibt die Vergleichbarkeit der Ergebnisse zwischen Ländern höchst umstritten.

Chancengleichheit

Das Konzept der Chancengleichheit gründet auf der normativen Ansicht, jeder Person unabhängig von sozialen, kulturellen und ökonomischen Bedingungen die gleichen Möglichkeiten zur Inanspruchnahme des Gesundheitssystems zu geben. Chancengleichheit bezieht sich in der Performance-Messung meist auf die Dimensionen des Zugangs bzw. der Nutzung von Gesundheitsleistungen sowie auf die Finanzierung des Gesundheitssystems.

Zentral bei der Chancengleichheit des Zugangs zum Gesundheitssystem ist die adäquate gesundheitliche Bedürfnisbefriedigung verschiedener gesellschaftlicher Gruppen. In diesem Zusammenhang wird zwischen horizontaler und vertikaler Dimension im Zugang unterschieden. Bei Ersterem handelt es sich um die Möglichkeit der Inanspruchnahme von Gesundheitsleistungen unterschiedlicher Gruppen bei äquivalenten Gesundheitsbedürfnissen. Dies wird häufig anhand von Indikatoren der Verfügbarkeit von Gesundheitsangeboten oder deren (Grad an) Nutzung beurteilt. Letzteres bezieht sich auf die Problematik, dass unterschiedliche Gesellschaftsgruppen differierende Bedürfnisse aufgrund von anders gelagerten Prioritäten und Werten bei gleicher gesundheitlicher Einschränkung haben

14 Bunker J. P. (2001): The role of medical care in contributing to health improvements within societies. *International Journal of Epidemiology*, 30(6).

15 Smith P.C., Mossialos E., Papanicolas I. (2008): op. cit., S. 3 u. 5.

16 WHO (2000), op. cit., 31f.





können. Die Verschiedenheit der Bedürfnisse suggeriert diversifizierte Gesundheitsangebote, um Patientenorientierung gewährleisten zu können. Allerdings ist die konzeptionelle Fassung der vertikalen Dimension in der Literatur – und damit die Ansatzpunkte für Messungen – stark umstritten.¹⁷ Finanzierung des Gesundheitssystems aus der Perspektive der Chancengleichheit basiert größtenteils auf der normativen Forderung von Zahlung nach Leistungsfähigkeit (ability to pay). Den Grund dafür sieht Wagstaff¹⁸ in der Wahllosigkeit von Individuen, entweder gesund oder krank zu sein. Konsum von Gesundheitsleistungen ist insofern keine freiwillige Entscheidung, sondern eine Notwendigkeit, um auf Krankheit adäquat zu reagieren. Dementsprechend sollte das (finanzielle) Risiko von Krankheit nicht nur von Individuen getragen und gleichzeitig die relative Distribution der verfügbaren Einkommen zumindest nicht verändert werden. Beispielsweise verwendete die WHO in ihrer Performance-Messung im Jahr 2000 einen Index, der die Güte der Finanzierung eines Gesundheitssystems anhand der Proportionalität der Zahlungen für alle Einkommensgruppen maß. Neben der Problematik, dass mit diesem Index nicht geklärt werden kann, ob vertikale Ungleichheit (z. B.: Personen mit höherem Einkommen leisten relativ höhere Gesundheitsausgaben) oder horizontale Ungleichheit (z. B.: Personen mit der gleichen Leistungsfähigkeit haben unterschiedlich hohe Gesundheitsausgaben) vorliegt,¹⁹ ist Proportionalität nur eine mögliche Variante der Auslegung des Leistungsfähigkeitsprinzips.

Effektivität

Effektivität ist eine wichtige Dimension der Performancemessung und beschreibt den Grad an Zielerreichung eines Gesundheitssystems bei gegebener Versorgung.²⁰ Die Zielerreichung ist abhängig von der politischen Entscheidung, welche Ziele in einem konkreten Gesundheitssystem erreicht werden sollen. Dies umfasst die bereits genannten Performance-Bereiche von Gesundheitszustand der Bevölkerung, Bedarfsgerechtigkeit und Chancengleichheit, aber auch andere Aspekte die als prioritär aufgefasst werden. Auch eine Standardisierung von Zielen über mehrere Länder hinweg, wie sie etwa durch die WHO postuliert wurde, ist möglich, um eine Vergleichbarkeit von Effektivität zu ermöglichen.

Effizienz

Effizienz kann als optimale Nutzung der Ressourcen zur Erbringung maximaler Ergebnisse definiert werden.²¹ Dabei wird zwischen technischer und allokativer Effizienz unterschieden. Mit technischer Effizienz wird entweder die Maximierung eines Outputs bei gegebenen Inputs oder die Minimierung des Inputs bei einem vorher festgelegten Output bezeichnet. Allokative Effizienz bezieht sich demgegenüber auf den optimalen Mix an Inputs bei gegebenen Input-Preisen, der die Kosten minimiert, oder auf den optimalen Mix an Outputs bei gegebenen Output-Preisen, der den Ertrag maximiert.²² Für die Performance-Messung kann Effizienz auf den verschiedenen Ebenen ansetzen. So kann man die effiziente Leistungserbringung von Ressourcen über das gesamte Gesundheitssystem messen, aber auch auf Anbieterebene bis hin zum einzelnen Patienten. Neben der schwierigen konzeptionellen Entscheidung über den Inhalt der Effizienzanalyse wird der buchhalterische Aufwand umso größer, je weiter man sich der Mikroebene nähert.²³

17 Gulliford M., Figueroa-Munoz J., Morgan M. (2003): Introduction: Meaning of "access" in health care, in: Gulliford M., Morgan M. (Ed.). Access to health care. London. Routledge. S. 8.

18 Wagstaff A. (2002): Reflections on an alternative to WHO's fairness of financial contribution

19 Wagstaff A. (2002), op.cit., S. 109.

20 Bankauskatie V., Dargent G. (2007): Health system performance indicators: methodological issues, S.127.

21 Joint Commission on Accreditation of Healthcare Organizations (1997): National Library of Health care Indicators. Health Plan and Network edition. Oakbrook Terrace, Illinois.

22 Hollingsworth, B. (2008): The measurement of efficiency and productivity of health care delivery, S. 1108.

23 Smith et al. (2008): op. cit.





Internationale Performance-Studien – einige Beispiele

World Health Report (WHR) 2000

Die Performance-Messung von Gesundheitssystemen im Zuge des World Health Report 2000 basiert für alle 191 WHO-Mitgliedsländer auf zwei wesentlichen Zielen von Gesundheitssystemen: Laut WHO gilt es einerseits, den bestmöglichen Level an Gesundheit (*goodness*) zu erreichen und andererseits die geringste gesundheitliche Differenz zwischen (Gruppen von) Menschen (*fairness*) sicherzustellen. Dabei konzeptionalisierte die WHO ein „Gesundheitssystem“ sehr breit und inkludierte alle Aktivitäten, die sich primär der Verbesserung, Wiederherstellung oder Beibehaltung des Gesundheitszustands widmen, in ihre Definition.²⁴ Zur Operationalisierung von „*goodness*“ und „*fairness*“ wurden fünf abgeleitete Ziele (*outcomes*) festgelegt, deren Erfüllungsgrad anhand von Indikatoren gemessen wurde. Diese sind:

- ◆ Maximierung des allgemeinen Gesundheitszustands, gemessen durch die Disability Adjusted Life Expectancy),
- ◆ höchstmögliche Gleichverteilung der Gesundheit in der Bevölkerung gemessen mit einem Inter-Individual-Differences-(IID-)Ansatz²⁵
- ◆ Maximierung des allgemeinen Niveau der Patienten-Bedürfnisentsprechung (Responsiveness) erhoben durch Fragebögen an „key informants“, beispielsweise zu den Bedürfnissen Würde, Autonomie oder Respekt,
- ◆ höchstmögliche Gleichverteilung der Patienten-Bedürfnisentsprechung (Responsiveness) erhoben anhand von Fragebögen an „key informants“ zu gruppenspezifischem Zugang zum Gesundheitssystem,
- ◆ höchstmögliche Gleichverteilung der Gesundheitskosten nach Maßgabe der Leistungsfähigkeit, bewertet nach dem Konzept „vertikaler Gleichheit“.²⁶

Die Ergebnisse flossen gewichtet in einen zusammengesetzten Indikator, der die Gesamtzielerreichung (*overall goal attainment*) des jeweiligen Landes darstellen soll. Um der unterschiedlichen Ausgangslage der einzelnen Länder an finanziellen und menschlichen Ressourcen Rechnung zu tragen, hat die WHO die Gesamtzielerreichung um diese Faktoren korrigiert. Das heißt, Länder werden nicht anhand ihrer absoluten Ergebnisse gemessen, sondern in Relation zu ihrem jeweiligem Potenzial²⁷ gesetzt. Der daraus berechnete zusammengesetzte Indikator der Gesamtperformance repräsentiert somit Gesundheitsniveau (*goal attainment*) und Effizienz der eingesetzten Mittel (*performance*).

Der WHR 2000 erfuhr große Aufmerksamkeit und in der Folge mannigfaltige Kritik auf verschiedenen Ebenen. Im Folgenden werden nun beispielhaft einige Hauptkritikpunkte dargestellt: *Erstens* standen für die Performance-Messung nur geringe Datenbestände zur Verfügung. 70 bis 89 Prozent (!), der in die Berechnungen zu den fünf Zielbereichen eingehenden Daten entstanden durch Imputationen. So wurde zum Beispiel die relative Verteilung der Gesundheitsausgaben auf die Einkommensklassen nur in 21 von 191 Ländern empirisch erhoben.²⁸ *Zweitens* basierte das Ranking der Patienten-Bedürfnisentsprechung auf Beantwortung von Fragebögen von zirka 1.800 „key informants“ aus nur 35 Ländern. Dabei waren diese Informanten primär Angestellte der WHO oder User der WHO-Website, die gebeten wurden einen Online-Fragebogen auszufüllen.²⁹ Zusätzlich wurden von einem Projektmitarbeiter Manipulationen am Datensatz gemeldet.³⁰ *Drittens* wurden die statis-

24 WHO (2000): op. cit., S. 5.

25 Dieses Maß beurteilt Ungleichverteilung von Gesundheit in der Bevölkerung anhand individueller Unterschiede und nicht, wie in der Literatur üblicher, nach gesellschaftlichen Gruppen (vgl. Asada/Hedemann: A problem with the individual approach in the WHO Health inequality measurement)

26 Je gleicher die relativen Anteile der Gesundheitsausgaben am Einkommen für alle Einkommensklassen, desto besser ist die Bewertung durch den World Health Report 2000.

27 Potenzial meint die optimale Nutzung vorhandener Ressourcen.

28 Almeida C., Braveman P., Gold M., Szwarcwald C. L., Mendes Ribeiro J., Miglionico A., Millar J., Porto S., Do Rosario Costa N., Ortun Rubio V., Segall M., Starfield B., Travassos C., Uga A., Valente J., Viacava F. (2001): Methodological concerns and recommendations on policy consequences of the World Health Report 2000. Lancet, 357, S. 1692. Nord E. (2002): Measures of goal attainment and performance in the World Health Report 2000: A brief, critical consumer guide. Health Policy, 59, S. 187.

29 Almeida C. et al. (2001): op. cit., S. 1693.

30 Musgrove P. (2003): Judging health systems: reflections on WHO's methods. Lancet, 361, S. 1818.

Internationale
Performance
Studien – einige
Beispiele





tischen Methoden für die Messung des Gesundheitszustands, die Verteilung der Gesundheit wie auch die Verteilung der Gesundheitskosten von einigen Kommentatoren als nicht valide eingestuft, wobei die Kritik von partieller Ausblendung einiger Aspekte bis zur Leugnung jeglicher Aussagekraft des Maßes reicht.³¹ *Viertens* verwendete der WHR beinahe ausschließlich Methoden, die WHO-intern entwickelt, jedoch keiner Veröffentlichung und damit keiner Expertenbeurteilung unterzogen wurden. Die Robustheit des statistischen Apparats war zum Zeitpunkt der Veröffentlichung völlig unklar.³² *Fünftens* kritisierten beispielsweise Coyne und Hilsenrath (2002),³³ dass es irreführend ist einen Effizienz- oder Systemindikator stark auf die Lebenserwartung zu stützen, wenn viele Determinanten der Lebenserwartung außerhalb des Gesundheitssektors liegen. Weiters wurde festgehalten, dass die WHO keinen adäquaten Versuch unternommen hat, zwischen Effizienz und Fairness zu unterscheiden. *Sechstens* werden auch die zuvor erwähnten geschätzten DALYs kritisiert. Reidparth et al. (2003)³⁴ halten fest, dass die DALYs für jede Region nur richtig sein können, wenn innerhalb jeder Region das Gewicht der Beeinträchtigung eines Gesundheitszustands dasselbe ist wie das fixe globale Gewicht der Beeinträchtigung, welches die WHO zur Schätzung der globalen Krankheitslast verwendet – dies kann nur der Fall sein, wenn das durchschnittliche soziale Milieu hinsichtlich eines Gesundheitszustands innerhalb einer Region dasselbe ist wie das globale durchschnittliche soziale Milieu, was natürlich einer völlig unwahrscheinlichen Situation entspricht.

European Health Consumer Index (EHCI)

Der im Jahre 2009 mittlerweile zum fünften Mal publizierte European Health Consumer Index (EHCI) umfasst die Analyse der Gesundheitssysteme von 33 europäischen Ländern. Ziel ist es, mit Hilfe von 38 ausgewählten Indikatoren die Gesundheitssysteme aus Konsumentensicht zu bewerten. Der Gesamtindex gliedert sich in sechs Bereiche, wobei sich jeder Bereich aus einer unterschiedlichen Anzahl von Indikatoren zusammensetzt und unterschiedlich gewichtet wird.

- ◆ Patientenrechte und -information (z. B. Recht auf zweite Meinung, Zugang zur eigenen Patientenakte, Vorhandensein eines Anbieterkatalogs mit Qualitätsranking,)
- ◆ E-Health (z. B. E-Datentransfer zwischen medizinischen Fachkräften, E-Rezept, Zugang zu Online-Terminvereinbarungen),
- ◆ Wartezeiten für Behandlungen (z. B. direkter Zugang zu Fachärzten, Krebstherapie in unter 21 Tagen),
- ◆ Behandlungsergebnisse (z. B. Säuglingssterblichkeit, Herzinfarkt-Sterblichkeit),
- ◆ Angebot und Reichweite der angebotenen Dienstleistungen (z. B. Mammographie-Screening, Nierentransplantationen pro Million Einwohner),
- ◆ Zugang zu Medikamenten (z. B. Zugang zu neuen Medikamenten, Kostenersatz für verschreibungspflichtige Medikamente).

Die Bewertung erfolgt durch eine dreistufige Skala, die die Ergebnisse als gut, mittelmäßig oder schlecht kategorisiert. Die erfolgte Kategorisierung der Indikatoren wird in Zahlen überführt, wobei sich der in die Messung eingehende numerische Wert, etwa für „gut“, nach der Gewichtung des bewerteten Indikators richtet. So werden für gute Ergebnisse bei Indikatoren im Bereich von Wartezeiten für Behandlungen doppelt so viele Punkte vergeben wie im Bereich Patientenrechte und -information.³⁵ Die Ergebnisse wurden zum Großteil aus öffentlich zugänglichen Informationen von nationalen Institutionen³⁶, internationale Organisationen (WHO, OECD, EU) oder eigenen Stu-

31 Almeida C. et al. (2001): op. cit., S. 1693 f. Ugá A. D., Almeida C., Szwarcwald C.L., Travessos C., Viacava F., Mendes Ribeiro J., do Rosario Costa N., Buss P. M., Porto S. (2001): Considerations on methodology used in the World Health Organization 2000 Report. Cad Saude Publica, 17(3), S. 708 ff. Richardson J., Wildman J., Robertson I. K. (2003): A critique of the World Health Organisation's evaluation of health system performance. Health Economics, 12, S. 358 ff.

32 Almeida C. et al. (2001): op. cit., S. 1695.

33 Coyne J.S., Hilsenrath P.: The World Health Report 2000: Can health care systems be compared using a single measure of performance?, American Journal of Public Health, 2002 January; 92(1): 30–33.

34 Reidparth D., Allety P. A., Kouame A., Cummins R. A.: Measuring health in a vacuum: examining the disability weight of the DALY, Health Policy and Planning, 2003 18(4): 351–356.

35 Björnberg A., Garrofé B. C., Lindblad S. (2009): Health Consumer Powerhouse: Euro Health Consumer Index Report. Brüssel, S. 27.

36 Keine nähere Definition vorhanden, ob Ministerien oder andere Behörden, Landesvertretungen oder nationale Forschungsinstitutionen, etc. befragt wurden.





dien zusammengetragen und laut den Autoren durch Experteninterviews, Behördenbesuche und Feedback-Meldungen von Behörden zu bereits vorab versendeten vorläufigen Studienergebnissen ergänzt.³⁷

Seit dem Beginn der Erstellung des EHCI 2005 haben sich der Umfang der untersuchten Länder, aber auch die verwendeten Indikatoren geändert. Die Ergebnisse der einzelnen Untersuchungen sind daher nur sehr eingeschränkt über die Zeit vergleichbar. Des Weiteren muss konstatiert werden, dass die Wissenschaftlichkeit des EHCI sehr zu wünschen übrig lässt: Quellen sind inhomogen und nicht transparent nachvollziehbar, und die willkürliche Gewichtung der Bereiche wurde über die Jahre durch die unterschiedliche Anzahl der verwendeten Indikatoren verändert. Auch ein Blick auf die einzelnen Bewertungen der Indikatoren lässt Fehleinschätzungen vermuten. Beispielsweise im Bereich der Wartezeiten, wo Österreich schlechter als Albanien abschneidet – zurückzuführen sicher darauf, dass sich die Indikatoren wohl auf unzureichend valide Datengrundlagen (in beiden Ländern) stützen. Die Gesundheit Österreich GmbH kritisiert zudem die Missinterpretation des Boxenmodells im Arzneimittelbereich – die Dauer bis zum Zugang zu neuen Medikamenten soll in Österreich angeblich über 300 Tage betragen. Abschließend muss generell die Eignung und die Sinnhaftigkeit einzelner Indikatoren zur Messung bestimmter Bereiche hinterfragt werden, zum Beispiel ob der Anteil der öffentlichen Gesundheitsausgaben an den gesamten Gesundheitsausgaben einen guten Indikator darstellt, die Fairness in einem Gesundheitssystem zu beurteilen.

Gesundheitsbarometer

Als mittlerweile vierte jährliche Studie analysiert das Gesundheitsbarometer 2010, durchgeführt vom Institut CSA (Cercle Santé Société) im Auftrag der zur Generali-Gruppe gehörenden Europ Assistance, die Einstellungen der Bevölkerungen zu Gesundheitsfragen in sieben europäischen Ländern (Österreich, Deutschland, Frankreich, Italien, Polen, Schweden, Vereinigtes Königreich) und in den Vereinigten Staaten.³⁸ Anders als die zwei zuvor beschriebenen Performance-Messungen findet beim Gesundheitsbarometer kein Ranking der Länder statt. Erhoben wurden ausschließlich Einstellungen von Konsumenten mit Hilfe von Telefoninterviews.³⁹ Diese wurden mit standardisierten Fragebögen durchgeführt, die entweder kategoriale oder ordinal-skalierte Antwortmöglichkeiten vorsahen. Insgesamt wurden 44 Fragen gestellt, die den folgenden Bereichen zugeordnet werden können:

- ◆ Zugang zum Gesundheitssystem und Qualität der Versorgung (z. B. Bewertung der Qualität von Behandlungen in Arztpraxen und Krankenhäusern),
- ◆ Fähigkeiten und Ausbildung von Ärzten (z. B. soziale und fachliche Fertigkeiten von Ärzten)
- ◆ soziale Bedürfnisse (z. B. bedarfsgerechter Zugang zu Gesundheitsleistungen für Patienten),
- ◆ Patientenmobilität im Bezug auf den Zugang von Behandlungen (z. B. Bereitschaft, Operationen im Ausland durchführen zu lassen),
- ◆ neue Technologien (z. B. Nutzung des Internets für Gesundheitsinformation seitens der Patienten),
- ◆ Prävention (z. B. Regelmäßigkeit der Inanspruchnahme von Vorsorgeuntersuchungen)
- ◆ Management von Pflege (z. B. Beurteilung der Versorgung von älteren und bedürftigen Menschen in den jeweiligen Ländern,)
- ◆ Selbst- und Fremdevaluierung von Gesundheitssystemen.

Die Problematiken der Gesundheitsbarometer-Studien liegen auf der Hand. Erfragt werden Meinungen von Nichtexperten, die die Konsumentenperspektive zu ausgewählten Themenbereichen darstellen können, jedoch nur beschränkte Aussagefähigkeit über die Effektivität und Effizienz des Gesundheitssystems zulassen. Einschätzungen über die Performance von Gesundheitssystemen sind durch vielerlei Faktoren (z. B. Medien, Ideologien) geprägt, die außerhalb des Gesundheitswe-

37 Björnberg A., Garrofé B. C., Lindblad S. (2009): op. cit, S. 42 ff.

38 Cercle Santé Société (2010): Healthcare in Europe and the USA. Presentation of the results, S. 2.

39 Insgesamt fanden 3600 Telefoninterviews statt, wobei 400 Personen über 18 Jahren in den jeweiligen europäischen Ländern und 800 in den USA befragt wurden. Dabei wurde auf Repräsentativität anhand der Kategorien Alter, Geschlecht, sozialer Status und Region geachtet.





sens liegen. Dies gilt insbesondere für die Einschätzungen von interviewten Personen über Gesundheitssysteme anderer Länder, die oftmals nicht einmal aus Konsumentenperspektive bekannt sind. Formal betrachtet ist auch die Veröffentlichung der Ergebnisse – ausschließlich in Präsentationsform ohne Langfassung in Berichtsform – als mangelhaft zu beurteilen. Schlussfolgerungen der Autoren aus den Ergebnissen sind daher unbekannt. Zudem existieren zwei verschiedene Präsentationsversionen der Studie für Deutschland und Österreich, ohne nachvollziehbaren Grund.⁴⁰

„Mirror, Mirror on the Wall“

„Mirror, Mirror on the Wall“ ist eine Studienreihe einer US-amerikanischen Privatstiftung namens „Commonwealth Fund“ zur Performancemessung ausgewählter nationaler Gesundheitswesen. Das 2010 veröffentlichte Ranking von sieben Gesundheitssystemen (Australien, Deutschland, Kanada, Neuseeland, Niederlande, Vereinigtes Königreich, Vereinigte Staaten) wurde anhand von 74 Indikatoren erstellt, deren Datengrundlagen im Rahmen von eigenen Studien zwischen 2007 und 2009 per Telefon⁴¹ erhoben wurden. Die verwendeten Datensätze sind in vier Dimensionen gegliedert und umfassen Patienten- und Ärzteperspektiven. Eine fünfte Kategorie „Allgemeine Gesundheit“ ergänzt die Studie durch Daten der OECD und WHO.

- ◆ Qualität (z. B. Indikatoren zu effektiver und patientenzentrierter Krankenpflege),
- ◆ Zugang (z. B. kostenbezogene Zugangsprobleme),
- ◆ Effizienz (z. B. Nutzung von Informationstechnologien oder Rehospitalisierungshäufigkeit),
- ◆ Gerechtigkeit (z. B. kostenbedingte Nichtinanspruchnahme eines Arztes)
- ◆ allgemeine Gesundheit (z. B. Säuglingssterblichkeit).

Die Bereichsrangings wie auch das Gesamtranking ergeben sich aus den gleichgewichteten Einzelergebnissen bzw. aus deren Rängen. Das Studiendesign basiert wie das Gesundheitsbarometer auf standardisierten Fragebögen mit dem Unterschied, dass neben der Konsumentenperspektive auch die Anbieterseite befragt wird. Zusätzlich fließen Daten zu allgemeinen Gesundheitsergebnissen in die Bewertung ein. Nichtsdestotrotz ist das Studiendesign für die komplexe Fragestellung der Performance nationaler Gesundheitswesen äußerst einfach konzipiert. Ähnlich wie bei allen bis jetzt vorgestellten Studien stellt sich die Frage der Validität der Ergebnisse insoweit, als soziale und ökonomische Bedingungen das Gesundheitswesen mitprägen, die Analyse hinsichtlich der Variablen aber unzureichend kontrolliert wird.

euroHOPE – European Health Outcomes, Performance and Efficiency

Die neueste (Groß-)Initiative in der Performance-Messung von Gesundheitssystemen in Europa ist das Projekt euroHOPE.⁴² Es wird im siebenten Rahmenprogramm der Europäischen Kommission (FP7) als vierjähriges Forschungsprogramm großteils von der EU finanziert und startete Anfang 2010. Getragen wird das Projekt von den finnischen Institutionen Center for Health and Social Economics (CHESS) und National Institute for Health and Welfare (THL). Ziel ist es, vorhandene Daten zu nutzen und neue Daten zu generieren, um die Performance von europäischen Gesundheitssystemen besonders in den Bereichen Herzinfarkt, Schlaganfall, Hüftfrakturen, Brustkrebs und Frühgeburten zu messen. Dabei werden folgende übergeordnete Ziele verfolgt:

- ◆ Entwicklung von Methoden zur vergleichenden Gesundheitssystemforschung,
- ◆ Entwicklung von Methoden zur Performance-Messung,
- ◆ Erstellung eines Indikatorenkatalogs für regelmäßige Performance-Messungen und Monitoring,
- ◆ Implementierung von europaweitem Benchmarking zur Nutzung von Best practice.

Es lohnt sich sicher, hier die weiteren Entwicklungen genau zu beobachten.

40 Cercle Santé Société (2010): op. cit. Online: http://www.europ-assistance.at/medias/pdf_intern/healthcare-in-europe-and-usa.pdf bzw. http://www.europ-assistance.de/medias/pdf_intern/100923_Gesundheitsbarometer_Studie.pdf

41 Insgesamt wurden ca. 27.500 Personen befragt, wobei 12.000 Personen ein repräsentatives Sample der Gesamtbevölkerungen über 18 Jahren, 8.700 Kranke und 6.800 niedergelassene Ärzte waren (Davis K., Schoen C., Stremikis K. (2010): Mirror, Mirror on the Wall. How the performance of the U.S. Health Care System Compares Internationally. New York, S. 21).

42 Vgl. <http://www.eurohope.info/>





Kritische Auseinandersetzung mit Gesundheitsrankings

Oft wird zur Bewertung eines Gesundheitssystems oder einer Teilkomponente davon ein zusammengesetztes Maß (Index) aus mehreren Bereichen berechnet, das in finaler Phase einer Rangfolge unterzogen wird. Dazu wird meist in einer Vorstufe je Bereich ein Index aus einzelnen Indikatoren gebildet. Zusammengesetzte Maße bieten den Vorteil, dass viele unterschiedliche Aspekte des Gesundheitssystems kompakt dargestellt werden können; durch die reduzierte Datenmenge ergibt sich für den Nutzer eine Erleichterung in Lesbarkeit und Interpretation. Jedoch bringt diese Vereinfachung auch große Gefahren mit sich. Einerseits ist es bei ausschließlicher Betrachtung des zusammengesetzten Maßes nicht mehr möglich, festzustellen, wo defizitäre Leistungsbereiche im Gesundheitssystem auftreten. Andererseits können durch die Nichtverwendung bestimmter Indikatoren Defizite in Teilbereichen des Gesundheitssystems „verschleiert“ werden. Ein zentraler Punkt bei der Berechnung von zusammengesetzten Maßen ist auch, inwiefern den einzelnen Indikatoren bzw. den einzelnen Bereichen eine Gewichtung gegeben wird. In manchen Untersuchungen erfolgt eine völlig willkürliche, nicht nachvollziehbare Gewichtung oder eine implizite Gewichtung durch die Verwendung einer unterschiedlichen Anzahl an Indikatoren je Bereich. Gewichtungen können einen Index stark in eine bestimmte Richtung lenken und sollten daher jedenfalls argumentierbar sein.

Jedes publizierte Gesundheits(system)-Ranking sollte einer genauen und kritischen Analyse unterzogen werden. Das Kanadische Institut für Gesundheitsinformation schlägt eine Checkliste für den Review von Rankings vor, vgl. Tabelle 2 und konzentriert sich dabei auf vier Kernfragen:⁴³

Tabelle 2: Checkliste für die Beurteilung von Studien mit Gesundheitsrankings

Checklist for Reviewing Health Ranking Reports

Step 1: Assess the soundness of the conceptual framework

- ☐ Does the ranking scheme's conceptual framework cover the areas of health and health care that are relevant to the purpose of the ranking?

Step 2: Assess the indicators chosen to measure selected aspects of health and health care

- ☐ Are the indicators of health or health care used in the ranking consistent with the conceptual framework?
- ☐ Are the measures used for the selected indicators meaningful and valid?

Step 3: Assess the data quality

- ☐ Are data accurate, reliable, complete, comparable and free from bias?
- ☐ Are data elements defined and collected so that „apples to apples“ comparisons are being made?

Step 4: Examine soundness of methods

- ☐ Are meaningful differences in performance distinguishable?
- ☐ Are absolute and relative comparisons available for review?
- ☐ Have appropriate adjustments been made for underlying differences in the populations being compared?
- ☐ Is the way specific measures are combined in the ranking scheme clear?
- ☐ Is the specific formula, along with any weights used to combine individual measures or indicators, based on clear and reasonable principles?
- ☐ Are differences in performance statistically significant?
- ☐ Have other statistical issues been appropriately handled (for example, adjustments for correlated measures or handling outlier values and ties)?
- ☐ Have the authors of the report reduced the potential for bias through full disclosure of ranking methods and peer review?

Quelle: Canadian Institute for Health Information: Making Sense of Health Rankings, Ottawa 2008.

Kritische Auseinandersetzung mit Gesundheitsrankings



43 Canadian Institute for Health Information (CIHI) (2008): op. cit., S. 17



1. Welche Aspekte von Gesundheit und des Gesundheitswesens werden im Ranking berücksichtigt?

Um die Aussagekraft einer Rankingstudie oder um Unterschiede zwischen derartigen Studien beurteilen zu können, muss in einem ersten Schritt evaluiert werden, ob der konzeptionelle Rahmen, auf den sich das Ranking stützt, gut fundiert ist und jene Gesundheitsbereiche ausreichend umfasst sind, die wichtig für die Zielsetzung des Rankings sind. Anschließend ist es notwendig, abzuschätzen, ob die für den Rankingprozess verwendeten Indikatoren konsistent mit dem konzeptionellen Rahmen sind.

Das Kanadische Institut für Gesundheitsinformation entwickelte beispielsweise einen konzeptionellen Rahmen, der drei große Bereiche abdeckt: Um den *Gesundheitszustand* einer Bevölkerung zu charakterisieren, sollte das gesundheitliche Wohlbefinden beurteilt oder auf Indikatoren zu Morbidität bzw. Mortalität zurückgegriffen werden. Zu den *nichtmedizinischen Determinanten von Gesundheit* zählen das Gesundheitsverhalten, Lebens- und Arbeitsbedingungen, Personalressourcen oder Umweltfaktoren. Die *Performance eines Gesundheitssystems* wird durch acht Dimensionen abgebildet: Akzeptanz, Zugänglichkeit, Angemessenheit, Kompetenz, Kontinuität, Effektivität, Effizienz und Sicherheit.

2. Wie aussagekräftig und valide sind die Indikatoren, die ausgewählt wurden, um diese Aspekte von Gesundheit und des Gesundheitswesens zu quantifizieren?

In einem zweiten Schritt muss beurteilt werden, ob für jeden Indikator aussagekräftige und valide Maße verwendet wurden. Fehlende Daten für einen oder mehrere Indikatoren sowie Probleme bei deren Messung erschweren den Einbezug wichtiger Aspekte für das Ranking. Entscheidend in diesem Zusammenhang ist, dass die Entwicklung und Standardisierung von Gesundheitsmaßen und neuen Datensammlungen sowohl auf nationaler als auch auf internationaler Ebene vorangetrieben wird. Im Zuge dessen würde sich auch die internationale Vergleichbarkeit der Indikatoren verbessern. Das Kanadische Institut für Gesundheitsinformation gibt einen Überblick, für welche Dimensionen zur Beurteilung der Performance eines Gesundheitssystems Indikatoren zur Verfügung stehen, vgl. Tabelle 3. Es ist zu erkennen, dass nicht für alle Dimensionen verfügbare Indikatoren vorhanden sind, und auch, dass einzelne Dimensionen sehr limitiert abgedeckt sind.

Tabelle 3: Beispiele für Gesundheitsindikatoren für die Bewertung der Performance des kanadischen Gesundheitssystems

Akzeptanz	Patientenzufriedenheit
Zugänglichkeit	Grippe-Immunisierung Mammographie-Screening Pap-Abstrich
Angemessenheit	Kaiserschnittgeburten Vaginale Geburten nach einem Kaiserschnitt
Effektivität	Keuchhusten-, Masern-, Tuberkulose-, HIV-, Chlamydieninfektionen Spitalsaufenthalte aufgrund von Lungenentzündung und Grippe Todesfälle aufgrund medizinisch behandelbarer Krankheiten (z. B.: bakterielle Infektionen, Gebärmutterhalskrebs, Hypertonie) 30-Tage-Herzinfarkt-Sterberate im Spital 30-Tage-Schlaganfall-Sterberate im Spital Wiederaufnahmeraten für Asthma, Gebärmutteroperationen, Lungenentzündungen, Prostataoperationen
Effizienz	Nicht benötigte Spitalsaufenthalte Erwarteter Aufenthalt verglichen mit tatsächlichem Aufenthalt
Sicherheit	Spitalsaufenthalte aufgrund einer Hüftfraktur Hüftfrakturen im Spital

Anmerkungen: Für Kompetenz und Kontinuität sind keine Maßzahlen verfügbar.
Quelle: Canadian Institute for Health Information 2008.

3. Basieren die verwendeten Indikatoren auf präzisen, zuverlässigen und vergleichbaren Informationen?

Im Idealfall sollten die verwendeten Daten aus Quellen stammen, die aktuell und vollständig sind sowie die Zielbevölkerung umfassen; Verzerrungen sollten so gut wie möglich ausgeschaltet werden. In diesem Kontext ist es besonders wichtig nachzuprüfen, wie einzelne Datenelemente definiert und erhoben wurden, um sicherzustellen, dass die Ergebnisse auch länderübergreifend vergleichbar sind. In den meisten Fällen wird bei internationalen Rankings auf offiziell publizierte Daten aus Datenban-





ken zurückgegriffen. Beispielsweise ist auch der häufig verwendete Indikator der Säuglingssterblichkeit mit Vorsicht zu genießen: Länder wie Kanada und die Vereinigten Staaten registrieren Todesfälle von Frühgeburten als Säuglingstodesfälle, andere Länder zählen diese Todesfälle zu den fötalen Sterbefällen; dieser Umstand könnte beispielsweise dazu beitragen, dass in den beiden genannten Ländern die Säuglingssterblichkeitsraten im Vergleich zu anderen Ländern erhöht sind.

4. Wie zuverlässig sind die Methoden, die dem Rankingprozess zugrunde liegen?

Performance-Rankings sollten den Anspruch haben, dass aussagekräftige Unterschiede in der Leistungsfähigkeit charakterisiert, absolute und relative Vergleiche interpretiert werden können sowie dass nachvollziehbar ist, wie die Indikatorenwerte im Rankingprozess kombiniert wurden. Zudem sollten statistisch signifikante Unterschiede in der Leistungsfähigkeit identifiziert werden können, und es soll darauf hingewiesen werden, welche Vor- bzw. Nachteile bestimmte Rankingverfahren aufweisen und wie sensitiv die Ergebnisse gegenüber der Verwendung anderer Methoden sind. Die Sensitivität der Ergebnisse einer internationalen OECD-Studie durch Verwendung unterschiedlicher Standardisierungs- und Rankingmethoden zeigte das IHS Ende 2009.⁴⁴

In der 2009 erschienenen Studie der OECD, „Doing Better for Children“,⁴⁵ wurde unter anderem das Wohlbefinden von Kindern im internationalen Kontext untersucht. In der vergleichenden Analyse von 30 OECD-Ländern erfolgte eine Darstellung von Rangfolgen für sechs verschiedene Dimensionen: Eine dieser Dimensionen bezog sich im engeren Sinne auf Gesundheit (und Sicherheit). Im Gegensatz zur 2007 veröffentlichten, ähnlich gestalteten Studie der UNICEF⁴⁶ verzichtete die OECD darauf, in ihrer Analyse ein Gesamtranking über alle Dimensionen hinweg darzustellen. Für den Bereich Gesundheit (und Sicherheit) wählte die OECD acht Indikatoren aus: niedriges Geburtsgewicht (in Prozent der Lebendgeburten), Säuglingssterblichkeit (Todesfälle pro 1.000 Lebendgeburten), Prozentsatz der Kinder, die jemals gestillt wurden, die beiden Immunisierungsraten gegen Keuchhusten und Masern (in Prozent), der Anteil der Kinder (11 bis 15 Jahre) die in der letzten Woche täglich einer mäßigen bis starken physischen Aktivität nachgegangen sind, sowie die Gesamtzahl an Todesfällen (unter 20 Jahren, pro 100.000 Einwohner) und die Anzahl der Todesfälle aufgrund von Suizid (15-19 Jahre, pro 100.000 Einwohner).

Das IHS modifizierte die Zusammensetzung der Indikatoren dahingehend, dass der Indikator „Niedriges Geburtsgewicht“ aufgrund einer sehr starken Korrelation mit den (Säuglings-)Mortalitätsraten ausgeschlossen wurde bzw. die Immunisierungsrate gegen Masern aus der Berechnung entfernt wurde. Die Gewichtung des Index in Richtung Säuglingsgesundheit wurde somit verringert und das Impfverhalten einer Bevölkerung nicht „doppelt“ abgebildet. Anschließend wurden einerseits genau die in der OECD-Studie verwendeten Methoden angewendet, was bedeutet, dass nach einheitlicher Polung der Indikatoren eine Standardisierung mit Hilfe einer Z-Transformation⁴⁷ und eine anschließende Mittelung durch den arithmetischen Mittelwert vollzogen wurden. Andererseits war es Ziel, durch Anwendung anderer Verfahren zu überprüfen, wie robust sich die Rankingergebnisse darstellen. Alternativ zur OECD-Methode erfolgte die Standardisierung der Indikatoren durch Bildung von Rangzahlen. Diese Art, eine Standardisierung vorzunehmen, bringt den Vorteil mit sich, dass für unterschiedlich starke Streuungen der Indikatoren sowie für Ausreißer korrigiert wird; umgekehrt zur Standardisierung auf Standardnormalverteilung, wo Indikatoren mit starken Streuungen und Extremwerten das Gesamtergebnis (bei Verwendung des arithmetischen Mittelwerts) auch stark in Richtung dieses Indikators lenken. Die Anwendung der Rangstandardisierung wird auch dadurch gerechtfertigt, dass die Streuung der hier verwendeten Indikatoren beträchtlich variiert.⁴⁸ Zusätzlich erfolgte im Zuge beider Standardisierungsverfahren eine anschließende Mittelung durch einen alternativen Lageparameter, nämlich durch den ausreißerrobusten Median.

Durch Verwendung der beiden Standardisierungsmethoden ergeben sich vier unterschiedliche Rangfolgen, die in Tabelle 4 dargestellt sind. Es zeigen sich zum Teil äußerst unterschiedliche Er-

44 Röhring G., Czipionka T., Riedel M.: Gesundheitsausgaben und -indikatoren im Kindesalter, Monatsschrift Kinderheilkunde Vol. 158(4), 04/2010, pp. 338–347 sowie IHS-Projektbericht im Auftrag des Hauptverbandes der österreichischen Sozialversicherungsträger (2009), unveröffentlicht.

45 OECD (2009): Doing Better for Children.

46 UNICEF (2007): An overview of child well-being in rich countries: A comprehensive assessment of the lives and well-being of children and adolescents in the economically advanced nations, Innocenti Report Card 7.

47 Standardisierung auf Standardnormalverteilung mit Mittelwert 0 und Varianz 1.

48 Variationskoeffizienten (Standardabweichung in Prozent des Mittelwerts) zwischen 8 und 94 Prozent.





Tabelle 4: Ergebnisse unterschiedlicher Berechnungsmethoden

	Ergebnisse der OECD-Studie	Z-Transformation – Arithmetischer Mittelwert Rangfolge 1	Z-Transformation – Median Rangfolge 2	Rangbildung – Arithmetischer Mittelwert Rangfolge 3	Rangbildung – Median Rangfolge 4	Arithmetischer Mittelwert aus Rangfolge 1 bis 4	Median aus Rangfolge 1 bis 4	Arithmetischer Mittelwert = Median	Finaler Rang
Anzahl der Indikatoren	8	6	6	6	6	6	6	6	6
Schweden	3	3	1	2	1	2	2	2	1
Island	2	2	2	1	2	2	2	2	2
Tschechien	5	4	8	3	3	5	4	4	3
Deutschland	9	7	5	4	4	5	5	5	4
Italien	17	5	6	6	8	6	6	6	5
Finnland	6	13	3	9	5	8	7	7	6
Japan	13	8	4	7	9	7	8	7	7
Luxemburg	7	6	10	5	12	8	8	8	8
Spanien	12	9	14	10	5	10	10	10	9
Dänemark	4	11	13	8	10	11	11	11	10
Slowakei	1	1	12	12	12	9	12	11	11
Niederlande	8	10	15	11	11	12	11	11	12
Frankreich	19	22	7	15	5	12	11	12	13
Ungarn	11	12	21	13	12	15	13	14	14
Schweiz	21	15	11	16	18	15	16	15	15
Griechenland	23	16	17	16	21	18	17	17	16
Portugal	18	17	9	18	17	15	17	16	17
Norwegen	16	21	20	14	16	18	18	18	18
Vereinigtes Königreich	20	14	23	19	21	19	20	20	19
Mexiko	28	29	18	22	15	21	20	21	20
Korea	10	18	19	24	24	21	22	21	21
Australien	15	19	16	26	27	22	23	22	22
Österreich	27	23	22	25	18	22	23	22	23
Polen	14	20	27	19	25	23	23	23	24
Irland	25	25	25	21	23	24	24	24	25
Türkei	30	28	26	22	20	24	24	24	26
Belgien	26	24	28	27	26	26	27	26	27
Kanada	22	26	24	28	28	27	27	27	28
USA	24	27	29	29	30	29	29	29	29
Neuseeland	29	30	30	30	29	30	30	30	30

Quelle: IHS HealthEcon Berechnungen 2010.

gebnisse: Die Slowakei beispielsweise rangiert bei Anwendung der OECD-Methode (Z-Transformation, arithmetischer Mittelwert) auf dem ersten Platz; Grund dafür ist, dass die Slowakei beim Indikator „physische Gesundheit“ einen extremen Ausreißer nach oben aufweist, der das Ergebnis des Gesamtrankings maßgeblich beeinflusst. Die Verwendung der anderen, nicht so stark ausreißer- und variationssensiblen Berechnungsmethoden illustriert äußerst robuste Ergebnisse für die Slowakei auf Rang 12. Das Ergebnis für Österreich verbessert sich leicht gegenüber jenem von der OECD publizierten; Österreich kommt jedoch deutlich hinter den anderen deutschsprachigen bzw. anderen vergleichbaren westeuropäischen Ländern zu liegen, und dies stabil.

Implikationen für die Politik

Implikationen für die Politik

Als ein finales Ziel von Performance-Messungen kann die Förderung des Erreichens von Zielen im Gesundheitssystem genannt werden. Die Messung von Leistung im Gesundheitswesen stellt aber nur ein Instrument dar, um Systemverbesserungen zu realisieren; erst in Kombination mit anderen Aspekten wie beispielsweise Finanzierung oder Regulierung sowie Aspekten, die Marktstruktur betreffend, können maximale Fortschritte erzielt werden. Den politischen Entscheidungsträgern kommt bei der Ausschöpfung des Potenzials der Performance-Messung die zentrale Aufgabe der Steuerung zu. Zum einen können Performance-Messungen den politisch Verantwortlichen helfen, Politiken zu formulieren, zu evaluieren oder als Regulativ einzugreifen; zum anderen erfüllt die Steuerungsfunktion auch die Sicherung eines notwendigen Informationsflusses und die Abstimmung auf die Ausgestaltung des Gesundheitssystems. Smith et al. op. cit. (2008) unterscheiden





zusammenfassend acht zentrale Steuerungsaufgaben in Verbindung mit Performance-Messung:

1. Im Fokus der *Entwicklung eines klaren konzeptionellen Rahmens und einer klaren Vision vom Zweck des Performance-Messungssystems*, soll die Abstimmung auf die Verantwortungsstruktur und auf andere Wirkmechanismen im Gesundheitssystem – Finanzierung, Marktstruktur, Informationstechnologie – erfolgen.
2. Die *Datenerhebung* sollte eine detaillierte Bestimmung der einzelnen Indikatoren beinhalten; zudem ist eine Abstimmung mit international anerkannten Praktiken der Datenerhebung notwendig.
3. Die *Datenpolitik* sollte dahingehend gestaltet sein, dass Datenaudit und Qualitätskontrolle stattfinden sowie eine Sicherung des öffentlichen Vertrauens in die Daten bzw. die Sicherung einer aufgeklärten öffentlichen Debatte gewährleistet ist.
4. Die Sicherung einer effizienten und effektiven Analyse sowie einer verständlichen Analyse für lokale Entscheidungsträger sollte in Verbindung mit Forschungsaufträgen in den Bereichen Risikoadjustierung oder Datenfeedback die *Entwicklung eines ausreichenden analytischen Apparats zum Verstehen der Daten* abstecken.
5. Durch die *Entwicklung geeigneter Verfahren zur Aggregation und Präsentation von Daten* sollte eine angemessene Wirkung und Vergleichbarkeit der Daten sowie die Verfügung der Veröffentlichung von Vergleichsdaten gesichert werden.
6. Leistungsmaße sollten aber auch *Verhaltensanreize* zur Folge haben, die einer Beobachtung unterzogen werden sollten.
7. Zentral ist auch eine *sachgerechte Bewertung der Leistungsmessungsinstrumente*, wie die Sicherung der kosteneffizienten Geldausgabe für Informationsquellen.
8. Der *politische Prozess* sollte dahingehend geleitet werden, dass Politikoptionen entwickelt und beobachtet werden sowie das Leistungsinformationssystem vor der Vereinnahmung durch einzelne Interessengruppen geschützt wird und schlussendlich eine gesunde politische Debatte initiiert und gefördert wird.

Fazit

Da bei jedem Gesundheits(system)-Ranking auf unterschiedliche Aspekte fokussiert wird und somit auch verschiedene Bereiche und Indikatoren verwendet werden, ist ein Vergleich von Rankings schier unmöglich. Es sind keine international vereinbarten Standards für gesundheitsbezogene Rankings etabliert. Dies erklärt u. a. auch, warum die Rangzahlen von Ländern bei unterschiedlichen Rankings oft einer starken Variation unterliegen. Insbesondere sind auch die Rankingmethoden zu hinterfragen und genau unter die Lupe zu nehmen. Eine ausschließliche Darstellung der Ergebnisse mit Hilfe einer Berechnungsmethode bzw. unzureichenden Hinweisen auf alternative Verfahren mit Vor- und Nachteilen ist unserer Ansicht nach nicht zulässig. Erschwert wird die Bewertung von Gesundheitssystemen durch die hochkomplexen Zusammenhänge, die in einfachen Indikatoren nur unzureichend abgebildet werden können. Unsere Beispiele aus einem Teilbereich des Gesundheitssystems, der Kindermedizin, zeigen bereits, wie anfällig Rankings auf Datenqualität und Indikatorenauswahl sind. Aussagen über die relative Güte ganzer Gesundheitssysteme sind deshalb mit besonderer Vorsicht zu genießen und dienen in erster Linie der politischen Vermarktung. Eine ernstzunehmende und wissenschaftlich fundierte Performance-Messung ist als Informationsinstrument aber ein anstrebenswertes Ziel. Sie wird sich aber nicht einfach auf Rankings herunterbrechen lassen, sondern eine höhere Informationsdichte und -komplexität besitzen.

Fazit





Health System Watch IV/2010

Übersicht internationaler Performance-Studien

	World Health Report	European Health Consumer Index	An overview of child well-being in rich countries (Dimension Gesundheit und Sicherheit)	Doing Better for Children (Dimension Gesundheit und Sicherheit)	Gesundheitsbarometer	Mirror, Mirror on the Wall	euroHOPE
Durchführung und Auftraggeber der Studie	World Health Organisation (WHO)	Health Consumer Powerhouse (HCP)	United Nations Children's Fund (UNICEF)	Organisation for Economic Co-operation and Development (OECD)	Institut CSA (Cercle Santé Société) im Auftrag der Europ Assistance Gruppe	Davis K., Schoen C., Stremikis K.; unterstützt durch den Commonwealth Fund	Center for Health and Social Economics (CHESS) und National Institute for Health and Welfare (THL) finanziert im Zuge des FP7 der Europäischen Kommission
Anzahl der bisher durchgeführten Studien	1	5	1	1	4	3	0
Jahr der letzten Veröffentlichung	2000	2009	2007	2009	2010	2010	Projektstart 2010
Anzahl der analysierten Länder	191	33	21	30	8	7	Start mit 7 Ländern; vorgesehen sind alle EU-Länder
Teilnahme Österreichs	ja	ja	ja	ja	ja	nein	ja, aber nicht bei den 7 Startländern
Gesamtranking/ Platzierung Österreichs	ja/9	ja/4	ja/18, Dimension Gesundheit und Sicherheit: 20	nein/Dimension Gesundheit und Sicherheit: 27	nein	ja/-	-
Bereiche (Dimensionen)	Gesundheitszustand	Patientenrechte und Information	materielles Wohlbefinden	materielles Wohlbefinden	Gesundheits-system	Zugang	Herzinfarkt
	Verteilung von Gesundheit	E-Health	Gesundheit und Sicherheit	Gesundheit und Sicherheit	Kompetenz Ausbildung Ärzte	Effizienz	Schlaganfall
	Responsiveness	Wartezeiten für Behandlungen	ausbildungsbezogenes Wohlbefinden	ausbildungsbezogenes Wohlbefinden	Soziale Bedürfnisse	Fairness	Hüftfrakturen
	Verteilung von Responsiveness	Gesundheits-ergebnisse	Risikoverhalten	Risikoverhalten	Patientenmobilität	Qualität	Brustkrebs
	Gesundheitskosten	Umfang der Versorgung	subjektives Wohlbefinden	Wohnen und Umwelt	neue Technologien	allgemeine Gesundheit	Frühgeburten
		Arzneimittel	Beziehungen zu Älteren und Gleichaltrigen	Qualität des Schullebens	Prävention		
					Ältere und bedürftige Menschen Selbst- und Fremdevaluierung von Gesundheitssystemen		
Anzahl der Indikatoren	11	38	40; Dimension Gesundheit und Sicherheit: 6	21; Dimension Gesundheit und Sicherheit: 8	44	74	noch nicht bekannt
Datengrundlagen	selbst erhobene Daten, key informants in 35 Ländern	selbst erhobene Daten und aus nationalen und intern. Datenbanken, Experteninterviews	Daten aus internationalen Datenbanken sowie Befragungsdaten	Daten aus internationalen Datenbanken sowie Befragungsdaten	Telefonbefragung	Telefonbefragung	noch nicht bekannt
Methode	gewichteter, zusammengesetzter Index	gewichteter, zusammengesetzter Index	ungewichteter Index je Dimension und für das Gesamtranking	ungewichteter Index je Dimension	Deskriptive Auswertung	gewichteter, zusammengesetzter Index	noch nicht bekannt
Kritik	konzeptionelle und methodische Schwächen, viele Imputationen	simplifiziertes Forschungsdesign, Recherchemängel	methodische Schwächen	methodische Schwächen	geringe Aussagekraft, Validitätsproblem	Forschungsdesign- und Validitätsproblem	noch zu früh für eine Bewertung

Quelle: IHS HealthEcon Darstellung 2010.

